REMARKS

Favorable reconsideration is respectfully requested in view of the foregoing amendments and the following remarks.

Pending Claims

Prior to this Amendment, Claims 1-17, 27, 30-34 and 40-43 were pending. All the pending claims have been cancelled and replaced by Claims 44 - 88.

Applicant reserves the right to file a divisional application with any of the claims cancelled herein.

Overview of claims

There are now 5 independent claims: 44, 47, 54, 58 and 68.

Claim 44 covers synthetic proteins that comprise retroinverted peptides with specified sequences.

Claim 47 covers synthetic proteins that comprise either retroinverted peptides with specified sequences or fragments of those retroinverted peptides.

Claim 54 covers synthetic proteins that comprise either retroinverted peptides with specified sequences or homologs of those retroinverted peptides.

Claim 58 covers synthetic proteins that comprise either retroinverted peptides with specified sequences, fragments of those retroinverted peptides, or homologs of those fragments.

Claim 68 covers synthetic proteins, up to 50 amino acids in length, that comprise retroinverted peptides with specified sequences. The specified sequences are shorter than the specified sequences in claims 47, 54, 58, and 58.

Incorporation by reference

Consistent with the guidelines in MPEP §608(p), Applicants are adding material from WO 98/51325 to the specification. WO 98/51325 is a published application that was incorporated by reference into the present application as can be seen from the following 2 excerpts from the present application:

> Previously, as disclosed and claimed in WO 98/51325, which is hereby incorporated by reference in its entirety, we have identified random peptides and their fragments, motifs, derivatives or peptidomimetics thereof which are capable of binding to GIT receptors such as the D2H, hSI, HPT1 and hPEPT1 receptors (hereinafter "GIT targeting peptides"). (From page 3, lines 7-11).

> The present invention relates to retro-inverted peptides (also referred to herein as "targeting retro-inverted peptides" or "targeting retro-inversion peptides") that target specific receptor sites in vivo and/or promote uptake of active agents and/or enhance active agent delivery across the GIT into the systemic, portal or hepatic circulation. In particular, these retro-inverted peptides specifically bind to one or more of the human gastrointestinal tract receptors HPT1, HPEPT1, D2H, or hSI or their equivalents in other mammals and have general utility in targeting active agents to selected sites and/or selected tissues in the body in which receptors are expressed. These peptides are synthesized from D-amino acids and have a reverse sequence order of the GIT targeting agents disclosed and claimed in the above-referenced WO 98/51325. (From page 4, line 20 to page 5, line12).

Material incorporated by reference from WO 98/51325 is summarized in the following table

Material	Location in WO 98/51325	Insert position in specification of present application	Claims in which Material appears in application
Information on GIT receptors	page 45 line 25 to page 46, line 37	Page 5, after line 11	None
Sequences of 55 receptor-binding peptides identified from a phage library (SEQ ID NOS: 16-70)	page 54, lines 5 to page 55, lines 37	Immediately following above insert	44, 47,54,58
Sequences of 13 binding motifs (SEQ ID NOS: 71-83)	Claims 6, 10, 14, 18-20	Sequence Listing	68
Sequences of 4 GIT receptors (SEQ ID NOS: 84-87)		Sequence Listing	None
Reference to 80 or 90 percent homology;	page 21, line 36 to page 22, line 16	Page 6, after line 14	54,56,58,66
fragment length is at least 5, 10 or 20 amino acids	page 21, line 36 to page 22, line 16	Page 6, after line 14	47-49,58-60
protein length is not more than 75 amino acids	page 21, line 36 to page 22, line 16	Page 6, after line 14	45,52,63

Changes made in text incorporated by reference

Applicants have incorporated text from page 54, line 5 to page 55, line 55 of WO 98/51325. The text corresponds to Table 7 of WO 98/51325 plus the paragraph that precedes it. Regarding that text, Applicants have made the following changes:

1) Added an introductory phrase to the paragraph preceding the Table: -- As indicated in WO 98/51325--

2) Added a sentence after the paragraph preceding the Table: -- Their insert sequences are summarized as follows: --

3) Deleted the header "Table 7"

4) Moved the title of the table, " TARGET BINDING PHAGE INSERT SEQUENCES" to become the header to the right column : --TARGET BINDING PHAGE INSERT SEQUENCE--

5) Changed the SEQ ID Nos from 1-55 to 16-70.

Support for Amendments

The following examples of support for any given claim limitation are intended to be illustrative, not exhaustive.

Support for newly added amino acid sequences

The SEQ ID NOs of newly added sequences incorporated by reference from WO 98/51325 are presented in the following Table together with their corresponding SEQ ID NOs from WO 98/51325.

SEQ ID NOs in present application	SEQ ID NOs in WO 98/51325	Nature of peptide/protein
16-70	1-55	Targeting agents
71-83	253-265	Targeting agents
84	176	hPEPT1 receptor
85	178	HPT1 receptor
86	179	hSI receptor
87	181	D2H receptor

Support for "specifically binds to a Caco-2 cell membrane fraction"

That phrase appears in the 5 newly added independent claims, 44, 47, 54, 58, and 68. The use of the Caco-2 assay to obtain data is described at pages 19-21. Regarding the Caco-2 assay, generally, as a test for the functionality of fragments and homologs, the following from the present application is noted:

> The present invention also relates to derivatives (including but not limited to fragments) of these retroinverted peptides, which derivatives are functionally similar to the retro-invert peptides (that is, capable of displaying one or more known functional activities of the retro-inverted peptides). These functional activities include but are not limited to the ability to bind or to compete with binding to the gastro-intestinal tract receptors HPT1, HPEPT1, D2H or hSI or the ability to be bound by an antibody directed against the retro-inverted peptide. Derivatives can be tested for the desired activity by procedures known in the art, including binding to a receptor domain or to Caco-2 cells, *in vitro*, or to intestinal tissue, *in vitro* or *in vivo*. (See page 5, lines 3-12, of the present application; underlining added here)

Support for the limitation that the synthetic protein does not exceed 75 amino acids in length

Support is found in material incorporated by reference from PCT application, page 21, line 36- page 22, line 5.

Support for the limitation that the synthetic protein does not exceed 50 amino acids in length

Support is found in Claim 4 of the present application as filed.

Support for the limitation that the fragments of specified retroinverted peptides are at least 5, 10 or 20 amino acids in length

Support is found in the material incorporated by reference from the PCT application, page 21, line 36 - page 22, line 4.

Support for the limitation that the homologs of specified retroinverted peptides show not more than 80 or 90 percent homology (but less than 100%)

Support for 80% and 90% is found in the material incorporated by reference from the PCT application, page 21, line 36 - page 22, line 11. Also, a "homolog", by definition, has less than 100% homology.

Support for the limitation that the homologs of specified retroinverted peptides meet one of four tests based of amino acid functional equivalency

This claim limitation, including the specification of 4 types of amino acid functional equivalency, finds support in the present application as filed, page 5, lines 24-29.

Support for claims which cover glycosylation, acetylation, phosphorylation, and amidation

Such claims find support in the present application as filed, page 5, line 30 to page 6, line 1.

Support for synthetic proteins with an added dansyl-lysine group

Such dansylated derivatives are made routinely for purposes of the CaCo-2 binding assay. (See pages 19-21 of the present application as filed).

Support for claims involving nanoparticles or microparticles, also size range

See claims 40-42 and pages 22-25 of the application as filed. As to particle sizes between 10 nm and 500 μm, see page 22, lines 5-8.

Support for drug classes and specific drugs covered in the Claims

See the application as filed, page 7, line 29 to page 9, line 1.

Support for the drug being insulin or leuprolide in the claims

See,the application as filed, claim 43 and pages 25-26.

Support for modifications to Table 1 of the present application

A number of changes have been made for clarity and consistency:

A column specifying the SEQ ID NO has been added at the left of the Table.

The K(dns) group has been eliminated from the sequence in rows 1 through 6. As a result, the sequences in rows 1-6 of the table now precisely reflect the sequences in the Sequence Listing previously submitted in this case for SEQ ID NOS: 1-6.

SEQ ID NOs 1-6, with their additonal K(dns) moieties, are now in rows 7-12 of Table 1. The K(dns) moiety is a dansyl-lysine mnoiety added to various peptides to make them detectable in the binding assays.

Modifications to Table 3 of the present application

Consistent with the amendments to Table 1, Table 3 has also been amended compared to the version submitted in the Amendment of October 5, 2001. The amendments are as follows:

Row 2, ZElan129, the SEQ ID NO: has been changed from 4 to 12.

Row 3, ZElan144, the SEQ ID NO: has been changed from 1 to 9.

Row 5, ZElan091, the SEQ ID NO: has been changed from 6 to 14.

Row 6, ZElan146, the SEQ ID NO: has been changed from 3 to 11.

Appendix to this Amendment

Applicants have attached an Applendix with copies of those pages from the WO 98/51325 that have the material that was incorporated via the present Amendment into the present application.

Sequence Listing

It is expected by the undersigned that an "AMENDMENT with Revised Sequence Listing" will be hand-delivered today to Group 1600 for Examiner Hope Robinson.

Response to rejections in Office Action of January 2, 2002.

Rejection of Claims 1-17, 27, 30-34 and 40-43 under 35 U.S.C. 112, first paragraph (Paragraph 2 of the Office Action)

The Examiner has rejected Claims 1-17, 27, 30-34 and 40-43 under 35 U.S.C. §112, first paragraph, stating that while being enabling for the retro-inverted peptides and the specific sequences (SEQ ID NOs: 1-3), the specification does not reasonably provide enablement for derivatives or fragments thereof or a binding portion thereof or a composition for treatment of any mammalian disease or disorder. This rejection is respectfully traversed for the reasons that follow. (Although the rejected claims having been replaced by the present Amendment, Applicants will respond to the rejection as if it was directed at each of the 5 independent claims now in the case.)

Claim **44** covers synthetic proteins that comprise retroinverted peptides with specified sequences. Data in the application shows examples where receptor binding ability is retained when an L-form peptide is "converted" to the retroinverted form.

Claim **47** covers synthetic proteins that comprise either retroinverted peptides with specified sequences or fragments of those retroinverted peptides. To the extent that some such fragments do not retain binding ability as specified in the claim, such fragments are not covered by the claim. To determine which fragments of a retroinverted peptide will retain that peptide's ability to bind in the Caco-2 binding assay, it is only necessary to identify the minimum "core region" needed for such binding. This can be done by systematically testing smaller and smaller fragments of a peptide for binding ability. In one approach, one successively eliminates 3-amino acid sections from each end of the 40-mer until binding ability is lost. If, for example, the core fragment is a 10-mer positioned at the center of the 40-mer, then the deletion of a 3-mer, 6-mer, 9-mer, 12-mer, and 15-mer from either end (10 tests total) would not eliminate the binding ability. Deletions of an 18-mer from either end would eliminate it. To achieve finer resolution, deletions of 16-mers and 17-mers could be tested. In any case, a total of only about 16 tests would be sufficient to identify the core binding peptide.

Claim **54** covers synthetic proteins that comprise either retroinverted peptides with specified sequences or homologs of those retroinverted peptides. As "homologs" implies similarity, the claim will tend to cover structures that retain binding ability. To the extent that some homologs do not retain the ability to bind as specified in the claim, such homologs are not covered by the claim.

Claim **58** covers synthetic proteins that comprise either retroinverted peptides with specified sequences, fragments of those retroinverted peptides, or homologs of those fragments. The fragments that retain specific binding activity can be determined in a reasonable number of steps as outlined above. As "homologs" implies similarity, the claim will tend to cover structures that retain binding ability. Fragment homologs that do not

retain the ability to bind are not covered by the claim.

Claim **68** covers synthetic proteins that comprise retroinverted peptides with specified sequences. Data in the application shows examples where receptor binding ability is retained when an L-form peptide is "converted" to the retroinverted form.

Applicants submit that the foregoing is responsive to the issues raised by the Examiner as to:

I. Quantitation of Experimentation;

II. Amount of direction or guidance presented;

IV. Nature of the invention;

V. State of the prior art and relative skill of those in the art; and

VI. Predictability or unpredictability of the art.

where the Roman numerals for each issue are those used by the Examiner.

The Examiner also raised issues **III** and **VII** as follows:

III. Presence or absence of working examples.

Applicants have included an example showing that orally delivered insulin-loaded nanoparticles coated with the retroinverted 15-mer peptide, ZElan144 produce as good or better bioavailability of insulin as such particles coated with ZElan 129, the L-peptide counterpart of ZElan 144 (Figure 2 and Table 5). The ZElan144-coated insulin-loaded nanoparticles also showed a therapeutic effect, evidenced by the reduction of glucose levels (Figure 1).

The retroinverted peptide ZElan 146 provided measureable bioavailability, about 20% that provided by ZElan 144.

Applicants submit that it is reasonable to extrapolate their success with ZElan 144 and ZElan 146 to the retroinverted forms of other peptides that are receptor binders.

VII. Breadth of the claims.

The Examiner has stated that the claims encompass any disease/disorder. In response, Applicants have amended the claims so that they are more specific as to the types of active agents envisioned. Applicants submit that, by providing more specificity as to what consititutes an active agent, Applicants inherently describe a corresponding disorder or disease known in the art to be treatable by that agent.

The Examiner has also stated that the claims cover any derivative/fragment or portion thereof. The claims presently in the case only cover those derivatives/fragments that show specific binding.

Rejection of Claims 1-17, 27, 30-34 and 40-43 under 35 U.S.C. 112, second paragraph (Paragraph 3 of the Office Action)

The Examiner has rejected Claims 1-17, 27, 30-34 and 40-43 under 35 U.S.C. §112, second paragraph, as being indefinite as follows:

1) Claim1 and dependent claims are rejected on the grounds that the recitation of "HPT1, hPEPT1, D2H and hSI" is insufficiently definite. Applicants no longer use these terms in the claims.

2) As to all the rejected claims, the Examiner has suggested using the qualifier "synthetic" or "isolated". Applicants use "synthetic" in the new claims.

3) Claims 2 and 13 are rejected on the grounds that "binding portion" is unclear. In the new claims, that term is not used.

4) Claims 4-7 are rejected on the grounds that they lack antecedent basis and suggests that Claim 1 be amended to recite specific sequences. The independent claims that have replaced Claim 1 recite specific sequences.

5) Claim 8 is rejected on the grounds that the meaning of the word "material" is

unclear. Although Claim 8 has been cancelled, the word "material" appears in new claims similar to Claim 8. In those claims (as in Claim 8), "Material" is intended to refer to any material that comprises the active agents specified in the claim, consistent with a major purpose of the invention - to be able to direct agent-loaded compositions to the GIT receptors.

6) Claims 8 and 13 are rejected on the grounds that no specific disease or disorder is described. As noted above, the new claims specify classes of drugs, and the drugs imply specific diseases.

7) Claim 16 is rejected on the grounds that it is unclear how the composition "facilitates" transport of the active agent. The word "facilitates" is not in the new claims.

8) Claim 30 is rejected on the grounds that that there is no antecedent basis for "one or more functional activities of said peptide". The phrase does not appear in the new claims.

In view of the foregoing remarks, it is respectfully submitted that all of the claims now pending in this application are allowable.

Respectfully submitted,
CAESAR, RIVISE, BERNSTEIN,
COHEN & POKOTILOW, LTD.

July 2, 2002

By Allan H. Fried

Allan H. Fried
Reg. No. 31,253
Seven Penn Center, 12th Floor
1635 Market Street
Philadelphia, PA 19103-2212
(215) 567-2010
Attorneys for Applicants

AMENDMENTS WITH MARKINGS SHOWING CHANGES

IN THE SPECIFICATION

Table 1, page 19, already amended on October 5, 2001, is further amended as follows:

--

SEQ ID NO:	Name	Description	Sequence
1	SEQ ID NO:1 [Zelan 144]	PAX2 15 mer fragment-D form retroinversion	rtrlrrnhsshkant [K(dns)-rtrlrrnhsshkant]
2	SEQ ID NO:2 [Zelan 145]	P31 16 mer fragment- D form retroinversion	gphrrgrpnsrsskrt [K(dns)- gphrrgrpnsrsskr]
3	SEQ ID NO:3 [Zelan 1146]	HAX42 14 mer fragment- D form retroinversion	gtsngngccnydgp [K(dns)- gtsngngccnydgp]
4	SEQ ID NO:4 [Zelan 129]	PAX2 15 mer fragment	TNAKHSSHNRRLRTR [K(dns)- TNAKHSSHNRRLRTR]
5	SEQ ID NO:5 [Zelan 031]	P31 16 mer fragment	TRKSSRSNPRGRRHPG [K(dns)- TRKSSRSNPRGRRHPG]
6	SEQ ID NO:6 [Zelan 091]	HAX42 14 mer fragment	PGDYNCCGNGNSTG [K(dns)- PGDYNCCGNGNSTG]
9	ZElan144	dansylated PAX2 15 mer fragment-D form retroinversion	K(dns)-rtrlrrnhsshkant
10	ZElan145	dansylated P31 16 mer fragment- D form retroinversion	K(dns)-gphrrgrpnsrsskrt
11	ZElan146	dansylated HAX42 14 mer fragment- D form retroinversion	K(dns)-gtsngngccnydgp
12	ZElan129	dansylated PAX2 15 mer fragment	K(dns)- TNAKHSSHNRRLRTR

SEQ ID NO:	Name	Description	Sequence
13	ZElan031	dansylated P31 16 mer fragment	K(dns)-TRKSSRSNPRGRRHPG
14	ZElan091	dansylated HAX42 14 mer fragment	K(dns)-PGDYNCCGNGNSTG

--

Table 3, page 21, already amended on October 5, 2001, is further amended as follows:

--

Name	Sequence	K_D (µmol)
ZElan018	K(dns)-STPPSREAYSRPYSVDSDSDTNAKHSSHNRRLRTRSRPNG (SEQ ID NO:7)	>50.0
ZElan129	K(dns)-TNAKHSSHNRRLRTR (SEQ ID [NO:4] NO:12)	29.6
ZElan144	K(dns)-rtrlrrnhsshkant (SEQ ID [NO:1] NO:9)	28.8
ZElan021	K(dns)-SDHALGTNLRSDNAKEPGDYNCCGNGNSTGRKVFNRRRPSAIPT (SEQ ID NO:8)	6.7
ZElan091	K(dns)-PGDYNCCGNGNSTG (SEQ ID [NO:6] NO:14)	0.75
ZElan146	K(dns)-gtsngngccnydgp (SEQ ID [NO:3] NO:11)	21.65

--

Please **replace** the paragraph at page 20, line 22 to page 21, line 2, already amended on October 5, 2001, with the following paragraph:

-- ZElan021, full length HAX42 [K(dns)-SDHALGTNLRSDNAKEPGDYNCCGNGNSTGRKVFNRRRPSAIPT] (SEQ ID NO:53; dansylated version is SEQ ID NO:8) was given the arbitrary value of 1.00 for binding to P100 at a given peptide concentration determined from the signal-to-noise ratio data. Table 2 shows the results of P100 assays with the HAX42 related peptides ZElan021, Zelan091 and ZElan146. Assay number 1 was at 20 μg/ml; 2 and 3 were at 50 μg/ml; and 4 through 8 were at 25 μg/ml. The results for the retro-inverted form, Zelan 146 show reasonable binding compared to the HAX42 fragment Zelan091 and that the activity of the GIT targeting agent was not eliminated when converted to its retro-inverted form. --

Please **replace** the paragraph at page 21, lines 5-11, already amended on October 5, 2001, with the following paragraph:

--K_D values, or the concentration of the peptide required to reach half maximal binding to Caco-2 P100 fractions, are given in Table 3 for ZElan021, full length HAX42, [K(dns)-SDHALGTNLRSDNAKEPGDYNCCGNGNSTGRKVFNRRRPSAIPT] (SEQ ID NO:53; dansylated version is SEQ ID NO:8), HAX42 fragment ZElan091, and the retro-inverted form of this fragment, ZElan146 as well as for ZElan018, full length PAX2, [K(dns)-STPPSREAYSRPYSVDS DSDTNAKHSSHNRRLRTRSRPNG] (SEQ ID NO:7; dansylated version is SEQ ID NO:15), PAX2 fragment ZElan129, and the retro-inverted form of this fragment, ZELan144.--

Appendix with pages from WO 98/512325

The following pages are attached:

21-22

45-46

54-55

179-180

184-189

192-194

234-237

Material incorporated by reference into the present application is marked by a vertical black line in the right margins.

known in the art, including binding to a GIT transport receptor domain or to Caco-2 cells, *in vitro*, or to intestinal tissue, *in vivo*. (See the Examples *infra*.)

In particular, derivatives can be made by altering GIT transport receptor-binding peptide sequences by substitutions, additions or deletions that provide for functionally equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other nucleotide sequences which encode substantially the same amino acid sequence may be used in the practice of the present invention. These include but are not limited to nucleotide sequences which are altered by the substitution of different codons that encode a functionally equivalent amino acid residue within the sequence, thus producing a silent change. Likewise, the GIT transport receptor-binding peptide derivatives of the invention include, but are not limited to, those containing, as a primary amino acid sequence, all or part of the amino acid sequence of a GIT transport receptor-binding peptide including altered sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence resulting in a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

In a specific embodiment of the invention, proteins consisting of or, alternatively, comprising all or a fragment

of a GIT transport receptor-binding peptide consisting of at least 5, 10, 15, 20, 25, 30 or 35 (contiguous) amino acids of the full-length GIT transport receptor-binding peptide are provided. In a specific embodiment, such proteins are not more than 20, 30, 40, 50, or 75 amino acids in length. Derivatives or analogs of GIT transport receptor-binding peptides include but are not limited to those molecules comprising regions that are substantially homologous to GIT transport receptor-binding peptides or fragments thereof (*e.g.*, at least 50%, 60%, 70%, 80% or 90% identity) (*e.g.*, over an identical size sequence or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art) or whose encoding nucleic acid is capable of hybridizing to a coding GIT transport receptor-binding peptide sequence, under stringent, moderately stringent, or nonstringent conditions.

In a specific embodiment, the GIT transport receptor-binding derivatives of the invention are not known proteins with homology to the GIT transport receptor-binding peptides of the invention or portions thereof.

The GIT transport receptor-binding peptide derivatives and analogs of the invention can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level. For example, the cloned GIT transport receptor-binding peptide gene sequence can be modified by any of numerous strategies known in the art (Maniatis, T., 1990, Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). The sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro*. In the production of the gene encoding a derivative or analog of GIT transport receptor-binding peptides, care should be taken to ensure that the modified gene remains within the same translational reading frame uninterrupted by translational

form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient.

The Therapeutics of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with free amino groups such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with free carboxyl groups such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

The amount of the Therapeutic of the invention which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances.

6. EXAMPLES

6.1. Selection of GIT Receptor Targets

The HPT1, hPEPT1, D2H, and hSI receptors were selected for cloning as GIT receptor targets based on several criteria, including: (1) expression on surface of epithelial cells in gastro-intestinal tract (GIT); (2) expression along the length of small intestine (HPT1, hPEPT1, D2H); (3) expression locally at high concentration (hSI); (4) large putative extracellular domains facing into the lumen of the GIT; and (5) extracellular domains that permit easy access and bioadhesion by targeting particles.

The four recombinant receptor sites screened with the peptide libraries additionally have the following characteristics:

Receptor	Characteristics
D2H	Transport of neutral/basic amino acids; a transport activating protein for a range of amino acid translocases
hSI	Metabolism of sucrose and other sugars; represents 9% of brush border membrane protein in Jejunum
HPT1	di/tri peptide transporter or facilitator of peptide transport
hPEPT1	di/tri peptide transporter

Figures 1-4 (SEQ ID NOS:176, 178, 179, and 181, respectively) show the predicted amino acid sequences for hPEPT1, HPT1, hSI and D2H, respectively.

6.2. Cloning of Extracellular Domain of Selected Receptor Site

The following receptor domains were cloned and expressed as His-tag fusion proteins by standard techniques:

Receptor	Domain (amino acid residues)
hPEPT1[a]	391-571
HPT1[b]	29-273
hSI[c]	272-667
D2H[d]	387-685

[a] Liang et al., 1995, J. Biol. Chem. 270:6456-6463
[b] Dantzig et al., 1994, Association of Intestinal Peptide Transport with a Protein Related to the Cadherin Superfamily
[c] Chantret et al., Biochem. J. 285:915-923
[d] Bertran et al., J. Biol. Chem. 268:14842-14949

The receptor proteins were expressed as His-tag fusion proteins and affinity purified under denaturing conditions, using urea or guanidine HCl, utilizing the pET His-tag metal chelate affinity for Ni-NTA Agarose (Hochuli, E., Purification of recombinant proteins with metal chelate adsorbent, Genetic Engineering, Principals and Methods (J.K. Setlow, ed.), Plenum Press, NY, Vol. 12 (1990), pp. 87-98).

plates were treated with PBS containing 0.1% phenylhydrazine for one hour at 37°C followed by two PBS washes and blocking for One hour with 0.5%BSA-PBS. The standard ELISA procedure was followed at this point.

Phage which showed specificity to a GIT receptor was further characterized by ELISA on a variety of recombinant proteins. Phage which continued to exhibit GIT receptor specificity was sequenced.

Table 7

TARGET BINDING PHAGE INSERT SEQUENCES:

hSI	SEQ. ID. NO.	
S15	1	RSGAYESPDGRGGRSYVGGGGGCGNIGRKHNLWGLRTASPACWD
S21	2	SPRSFWPVVSRHESFGISNYLGCGYRTCISGTMTKSSPIYPRHS
S22	3	SSSSDWGGVPGKVVRERFKGRGCGISITSVLTGKPNPCPEPKAA
SNi10	4	RVGQCTDSDVRRPWARSCAHQGCGAGTRNSHGCITRPLRQASAH
SNi28	5	SHSGGMNRAYGDVFRELRDRWNATSHHTRPTPQLPRGPN
SNi34	6	SPCGGSWGRFMQGGLFGGRTDGCGAHRNRTSASLEPPSSDY
SNi38	7	RGAADQRRGWSENLGLPRVGWDAIAHNSYTFTSRRPRPP
SNi45	8	SGGEVSSWGRVNDLCARVSWTGCGTARSARTDNKGFLPKHSSLR
SNiAX2	9	SDSDGDHYGLRGGVRCSLRDRGCGLALSTVHAGPPSFYPKLSSP
SNiAX4	10	RSLGNYGVTGTVDVTVLPMPGHANHLGVSSASSSDPPRR
SNiAX6	11	RTTTAKGCLLGSFGVLSGCSFTPTSPPPHLGYPPHSVN
SNiAX8	12	SPKLSSVGVMTKVTELPTEGPNAISIPISATLGPRNPLR
D2H		
DAB3	13	RWCGAELCNSVTKKFRPGWRDHANPSTHHRTPPPSQSSP
DAB7	14	RWCGADDPCGASRWRGGNSLFGCGLRCSAAQSTPSGRIHSTSTS
DAB10	15	SKSGEGGDSSRGETGWARVRSHAMTAGRFRWYNQLPSDR
DAB18	16	RSSANNCEWKSDWMRRACIARYANSSGPARAVDTKAAP
DAB24	17	SKWSWSSRWGSPQDKVEKTRAGCGGSPSSTNCHPYTFAPPPQAG
DAB30	18	SGFWEFSRGLWDGENRKSVRSGCGFRGSSAQGPCPVTPATIDKH
DAX15	19	SESGRCRSVSRWMTTWQTQKGGCGSNVSRGSPLDPSHQTGHATT
DAX23	20	REWRFAGPPLDLWAGPSLPSFNASSHPRALRTYWSQRPR
DAX24	21	RMEDIKNSGWRDSCRWGDLRPGCGSRQWYPSNMRSSRDYPAGGH
DAX27	22	SHPWYRHWNHGDFSGSGQSRHTPPESPHPGRPNATI

DCX8	23	RYKHDIGCDAGVDKKSSSVRGGCGAHSSPPRAGRGPRGTMVSRL
DCX11	24	SQGSKQCMQYRTGRLTVGSEYGCGMNPARHATPAYPARLLPRYR
DCX26	25	SGRTTSEISGLWGWGDDRSGYGWGNTLRPNYIPYRQATNRHRYT
DCX33	26	RWNWTVLPATGGHYWTRSTDYHAINNHRPSIPHQHPTPI
DCX36	27	SWSSWNWSSKTTRLGDRATREGCGPSQSDGCPYNGRLTTVKPRT
DCX39	28	SGSLNAWQPRSWVGGAFRSHANNNLNPKPTMVTRHPT
DCX42	29	RYSGLSPRDNGPACSQEATLEGCGAQRLMSTRRKGRNSRPGWTL
DCX45	30	SVGNDKTSRPVSFYGRVSDLWNASLMPKRTPSSKRHDDG
<u>hPEPT1</u>		
PAX9	31	RWPSVGYKGNGSDTIDVHSNDASTKRSLIYNHRRPLFP
PAX14	32	RTFENDGLGVGRSIQKKSDRWYASHNIRSHFASMSPAGK
PAX15	33	SYCRVKGGGEGGHTDSNLARSGCGKVARTSRLQHINPRATPPSR
PAX16	34	SWTRWGKHTHGGFVNKSPPGKNATSPYTDAQLPSDQGPP
PAX17	35	SQVDSFRNSFRWYEPSRALCHGCGKRDTSTTRIHNSPSDSYPTR
PAX18	36	SFLRFQSPRFEDYSRTISRLRNATNPSNVSDAHNNRALA
PAX35	37	RSITDGGINEVDLSSVSNVLENANSHRAYRKHRPTLKRP
PAX38	38	SSKVSSPRDPTVPRKGGNVDYGCGHRSSARMPTSALSSITKCYT
PAX40	39	RASTQGGRGVAPEFGASVLGRGCGSATYYTNSTSCKDAMGHNYS
PAX43	40	RWCEKHKFTAARCSAGAGFERDASRPPQPAHRDNTNRNA
PAX45	41	SFQVYPDHGLERHALDGTGPLYAMPGRWIRARPQNRDRQ
PAX46	42	SRCTDNEQCPDTGTRSRSVSNARYFSSRLLKTHAPHRP
P31	43	SARDSGPAEDGSRAVRLNGVENANTRKSSRSNPRGRRHP
P90	44	SSADAEKCAGSLLWWGRQNNSGCGSPTKKHLKHRNRSQTSSSSH
5PAX3	45	RPKNVADAYSSQDGAAAEETSHASNAARKSPKHKPLRRP
5PAX5	46	RGSTGTAGGERSGVLNLHTRDNASGSGFKPWYPSNRGHK
5PAX7	47	RWGWERSPSDYDSDMDLGARRYATRTHRAPPRVLKAPLP
5PAX12	48	RGWKCEGSQAAYGDKDIGRSRGCGSITKNNTNHAHPSHGAVAKI
<u>HPT-1</u>		
HAX9	49	SREEANWDGYKREMSHRSRFWDATHLSRPRRPANSGDPN
HAX35	50	EWYSWKRSSKSTGLGDTATREGCGPSQSDGCPYNGRLTTVKPRK
HAX40	51	REFAERRLWGCDDLSWRLDAEGCGPTPSNRAVKHRKPRPRSPAL
HAX42	52	SDHALGTNLRSDNAKEPGDYNCCGNGNSTGRKVFNRRRPSAIPT
HCA3	53	RHISEYSFANSHLMGGESKRKGCGINGSFSPTCPRSPTPAFRRT
H40	54	SRESGMWGSWWRGHRLNSTGGNANMNASLPPDPPVSTP
PAX2	55	STPPSREAYSRPYSVDSDSDTNAKHSSHNRRLRTRSRPN

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

Arg Arg Leu Arg Thr Arg Ser Arg Pro Asn
1 5 10

(2) INFORMATION FOR SEQ ID NO:176:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 708 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

Met Gly Met Ser Lys Ser His Ser Phe Phe Gly Tyr Pro Leu Ser Ile
1 5 10 15
Phe Phe Ile Val Val Asn Glu Phe Cys Glu Arg Phe Ser Tyr Tyr Gly
20 25 30
Met Arg Ala Ile Leu Ile Leu Tyr Phe Thr Asn Phe Ile Ser Trp Asp
35 40 45
Asp Asn Leu Ser Thr Ala Ile Tyr His Thr Phe Val Ala Leu Cys Tyr
50 55 60
Leu Thr Pro Ile Leu Gly Ala Leu Ile Ala Asp Ser Trp Leu Gly Lys
65 70 75 80
Phe Lys Thr Ile Val Ser Leu Ser Ile Val Tyr Thr Ile Gly Gln Ala
85 90 95
Val Thr Ser Val Ser Ser Ile Asn Asp Leu Thr Asp His Asn His Asp
100 105 110
Gly Thr Pro Asp Ser Leu Pro Val His Val Val Leu Ser Leu Ile Gly
115 120 125
Leu Ala Leu Ile Ala Leu Gly Thr Gly Gly Ile Lys Pro Cys Val Ser
130 135 140
Ala Phe Gly Gly Asp Gln Phe Glu Glu Gly Gln Glu Lys Gln Arg Asn
145 150 155 160
Arg Phe Phe Ser Ile Phe Tyr Leu Ala Ile Asn Ala Gly Ser Leu Leu
165 170 175
Ser Thr Ile Ile Thr Pro Met Leu Arg Val Gln Gln Cys Gly Ile His
180 185 190
Ser Lys Gln Ala Cys Tyr Pro Leu Ala Phe Gly Val Pro Ala Ala Leu
195 200 205
Met Ala Val Ala Leu Ile Val Phe Val Leu Gly Ser Gly Met Tyr Lys
210 215 220
Lys Phe Lys Pro Gln Gly Asn Ile Met Gly Lys Val Ala Lys Cys Ile
225 230 235 240
Gly Phe Ala Ile Lys Asn Arg Phe Arg His Arg Ser Lys Ala Phe Pro
245 250 255
Lys Arg Glu His Trp Leu Asp Trp Ala Lys Glu Lys Tyr Asp Glu Arg
260 265 270
Leu Ile Ser Gln Ile Lys Met Val Thr Arg Val Met Phe Leu Tyr Ile
275 280 285
Pro Leu Pro Met Phe Trp Ala Leu Phe Asp Gln Gln Gly Ser Arg Trp
290 295 300
Thr Leu Gln Ala Thr Thr Met Ser Gly Lys Ile Gly Ala Leu Glu Ile
305 310 315 320
Gln Pro Asp Gln Met Gln Thr Val Asn Ala Ile Leu Ile Val Ile Met

```
                    325                 330                 335
Val Pro Ile Phe Asp Ala Val Leu Tyr Pro Leu Ile Ala Lys Cys Gly
            340                 345                 350
Phe Asn Phe Thr Ser Leu Lys Lys Met Ala Val Gly Met Val Leu Ala
        355                 360                 365
Ser Met Ala Phe Val Val Ala Ala Ile Val Gln Val Glu Ile Asp Lys
    370                 375                 380
Thr Leu Pro Val Phe Pro Lys Gly Asn Glu Val Gln Ile Lys Val Leu
385                 390                 395                 400
Asn Ile Gly Asn Asn Thr Met Asn Ile Ser Leu Pro Gly Glu Met Val
                405                 410                 415
Thr Leu Gly Pro Met Ser Gln Thr Asn Ala Phe Met Thr Phe Asp Val
            420                 425                 430
Asn Lys Leu Thr Arg Ile Asn Ile Ser Ser Pro Gly Ser Pro Val Thr
        435                 440                 445
Ala Val Thr Asp Asp Phe Lys Gln Gly Gln Arg His Thr Leu Leu Val
    450                 455                 460
Trp Ala Pro Asn His Tyr Gln Val Val Lys Asp Gly Leu Asn Gln Lys
465                 470                 475                 480
Pro Glu Lys Gly Glu Asn Gly Ile Arg Phe Val Asn Thr Phe Asn Glu
                485                 490                 495
Leu Ile Thr Ile Thr Met Ser Gly Lys Val Tyr Ala Asn Ile Ser Ser
            500                 505                 510
Tyr Asn Ala Ser Thr Tyr Gln Phe Phe Pro Ser Gly Ile Lys Gly Phe
        515                 520                 525
Thr Ile Ser Ser Thr Glu Ile Pro Pro Gln Cys Gln Pro Asn Phe Asn
    530                 535                 540
Thr Phe Tyr Leu Glu Phe Gly Ser Ala Tyr Thr Tyr Ile Val Gln Arg
545                 550                 555                 560
Lys Asn Asp Ser Cys Pro Glu Val Lys Val Phe Glu Asp Ile Ser Ala
                565                 570                 575
Asn Thr Val Asn Met Ala Leu Gln Ile Pro Gln Tyr Phe Leu Leu Thr
            580                 585                 590
Cys Gly Glu Val Val Phe Ser Val Thr Gly Leu Glu Phe Ser Tyr Ser
        595                 600                 605
Gln Ala Pro Ser Asn Met Lys Ser Val Leu Gln Ala Gly Trp Leu Leu
    610                 615                 620
Thr Val Ala Val Gly Asn Ile Ile Val Leu Ile Val Ala Gly Ala Gly
625                 630                 635                 640
Gln Phe Ser Lys Gln Trp Ala Glu Tyr Ile Leu Phe Ala Ala Leu Leu
                645                 650                 655
Leu Val Val Cys Val Val Phe Ala Ile Met Ala Arg Phe Tyr Thr Tyr
            660                 665                 670
Ile Asn Pro Ala Glu Ile Glu Ala Gln Phe Asp Glu Asp Glu Lys Lys
        675                 680                 685
Asn Arg Leu Glu Lys Ser Asn Pro Tyr Phe Met Ser Gly Ala Asn Ser
    690                 695                 700
Gln Lys Gln Met
705
```

(2) INFORMATION FOR SEQ ID NO:177:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3345 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA
(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 88...2583
(D) OTHER INFORMATION:

```
CAC CAG ACT GGG ATA CCC ACT GTG GGC ATG GCA GTT GGT ATA CTG CTG     2466
His Gln Thr Gly Ile Pro Thr Val Gly Met Ala Val Gly Ile Leu Leu
        780                 785                 790

ACC ACC CTT CTG GTG ATT GGT ATA ATT TTA GCA GTT GTG TTT ATC CGC     2514
Thr Thr Leu Leu Val Ile Gly Ile Ile Leu Ala Val Val Phe Ile Arg
    795                 800                 805

ATA AAG AAG GAT AAA GGC AAA GAT AAT GTT GAA AGT GCT CAA GCA TCT     2562
Ile Lys Lys Asp Lys Gly Lys Asp Asn Val Glu Ser Ala Gln Ala Ser
810                 815                 820                 825

GAA GTC AAA CCT CTG AGA AGC TGAATTTGAA AAGGAATGTT TGAATTTATA TAGC   2617
Glu Val Lys Pro Leu Arg Ser
                830

AAGTGCTATT TCAGCAACAA CCATCTCATC CTATTACTTT TCATCTAACG TGCATTATAA    2677
TTTTTTAAAC AGATATTCCC TCTTGTCCTT TAATATTTGC TAAATATTTC TTTTTTGAGG    2737
TGGAGTCTTG CTCTGTCGCC CAGGCTGGAG TACAGTGGTG TGATCCCAGC TCACTGCAAC    2797
CTCCGCCTCC TGGGTTCACA TGATTCTCCT GCCTCAGCTT CCTAAGTAGC TGGGTTTACA    2857
GGCACCCACC ACCATGCCCA GCTAATTTTT GTATTTTTAA TAGAGACGGG GTTTCGCCAT    2917
TTGGCCAGGC TGGTCTTGAA CTCCTGACGT CAAGTGATCT GCCTGCCTTG GTCTCCCAAT    2977
ACAGGCATGA ACCACTGCAC CCACCTACTT AGATATTTCA TGTGCTATAG ACATTAGAGA    3037
GATTTTTCAT TTTTCCATGA CATTTTTCCT CTCTGCAAAT GGCTTAGCTA CTTGTGTTTT    3097
TCCCTTTTGG GGCAAGACAG ACTCATTAAA TATTCTGTAC ATTTTTTCTT TATCAAGGAG    3157
ATATATCAGT GTTGTCTCAT AGAACTGCCT GGATTCCATT TATGTTTTTT CTGATTCCAT    3217
CCTGTGTCCC CTTCATCCTT GACTCCTTTG GTATTTCACT GAATTTCAAA CATTTGTCAG    3277
AGAAGAAAAA AGTGAGGACT CAGGAAAAAT AAATAAATAA AAGAACAGCC TTTTGCGGCC    3337
GCGAATTC                                                            3345
```

(2) INFORMATION FOR SEQ ID NO:178:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 832 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

```
Met Ile Leu Gln Ala His Leu His Ser Leu Cys Leu Leu Met Leu Tyr
 1               5                  10                  15
Leu Ala Thr Gly Tyr Gly Gln Glu Gly Lys Phe Ser Gly Pro Leu Lys
            20                  25                  30
Pro Met Thr Phe Ser Ile Tyr Glu Gly Gln Glu Pro Ser Gln Ile Ile
        35                  40                  45
Phe Gln Phe Lys Ala Asn Pro Pro Ala Val Thr Phe Glu Leu Thr Gly
    50                  55                  60
Glu Thr Asp Asn Ile Phe Val Ile Glu Arg Glu Gly Leu Leu Tyr Tyr
65                  70                  75                  80
Asn Arg Ala Leu Asp Arg Glu Thr Arg Ser Thr His Asn Leu Gln Val
                85                  90                  95
Ala Ala Leu Asp Ala Asn Gly Ile Ile Val Glu Gly Pro Val Pro Ile
            100                 105                 110
Thr Ile Glu Val Lys Asp Ile Asn Asp Asn Arg Pro Thr Phe Leu Gln
        115                 120                 125
Ser Lys Tyr Glu Gly Ser Val Arg Gln Asn Ser Arg Pro Gly Lys Pro
    130                 135                 140
Phe Leu Tyr Val Asn Ala Thr Asp Leu Asp Asp Pro Ala Thr Pro Asn
145                 150                 155                 160
Gly Gln Leu Tyr Tyr Gln Ile Val Ile Gln Leu Pro Met Ile Asn Asn
                165                 170                 175
Val Met Tyr Phe Gln Ile Asn Asn Lys Thr Gly Ala Ile Ser Leu Thr
            180                 185                 190
```

```
Arg Glu Gly Ser Gln Glu Leu Asn Pro Ala Lys Asn Pro Ser Tyr Asn
        195                 200                 205
Leu Val Ile Ser Val Lys Asp Met Gly Gly Gln Ser Glu Asn Ser Phe
    210                 215                 220
Ser Asp Thr Thr Ser Val Asp Ile Ile Val Thr Glu Asn Ile Trp Lys
225                 230                 235                 240
Ala Pro Lys Pro Val Glu Met Val Glu Asn Ser Thr Asp Pro His Pro
                245                 250                 255
Ile Lys Ile Thr Gln Val Arg Trp Asn Asp Pro Gly Ala Gln Tyr Ser
            260                 265                 270
Leu Val Asp Lys Glu Lys Leu Pro Arg Phe Pro Phe Ser Ile Asp Gln
        275                 280                 285
Glu Gly Asp Ile Tyr Val Thr Gln Pro Leu Asp Arg Glu Glu Lys Asp
    290                 295                 300
Ala Tyr Val Phe Tyr Ala Val Ala Lys Asp Glu Tyr Gly Lys Pro Leu
305                 310                 315                 320
Ser Tyr Pro Leu Glu Ile His Val Lys Val Lys Asp Ile Asn Asp Asn
                325                 330                 335
Pro Pro Thr Cys Pro Ser Pro Val Thr Val Phe Glu Val Gln Glu Asn
            340                 345                 350
Glu Arg Leu Gly Asn Ser Ile Gly Thr Leu Thr Ala His Asp Arg Asp
        355                 360                 365
Glu Glu Asn Thr Ala Asn Ser Phe Leu Asn Tyr Arg Ile Val Glu Gln
    370                 375                 380
Thr Pro Lys Leu Pro Met Asp Gly Leu Phe Leu Ile Gln Thr Tyr Ala
385                 390                 395                 400
Gly Met Leu Gln Leu Ala Lys Gln Ser Leu Lys Lys Gln Asp Thr Pro
                405                 410                 415
Gln Tyr Asn Leu Thr Ile Glu Val Ser Asp Lys Asp Phe Lys Thr Leu
            420                 425                 430
Cys Phe Val Gln Ile Asn Val Ile Asp Ile Asn Asp Gln Ile Pro Ile
        435                 440                 445
Phe Glu Lys Ser Asp Tyr Gly Asn Leu Thr Leu Ala Glu Asp Thr Asn
    450                 455                 460
Ile Gly Ser Thr Ile Leu Thr Ile Gln Ala Thr Asp Ala Asp Glu Pro
465                 470                 475                 480
Phe Thr Gly Ser Ser Lys Ile Leu Tyr His Ile Ile Lys Gly Asp Ser
                485                 490                 495
Glu Gly Arg Leu Gly Val Asp Thr Asp Pro His Thr Asn Thr Gly Tyr
            500                 505                 510
Val Ile Ile Lys Lys Pro Leu Asp Phe Glu Thr Ala Ala Val Ser Asn
        515                 520                 525
Ile Val Phe Lys Ala Glu Asn Pro Glu Pro Leu Val Phe Gly Val Lys
    530                 535                 540
Tyr Asn Ala Ser Ser Phe Ala Lys Phe Thr Leu Ile Val Thr Asp Val
545                 550                 555                 560
Asn Glu Ala Pro Gln Phe Ser Gln His Val Phe Gln Ala Lys Val Ser
                565                 570                 575
Glu Asp Val Ala Ile Gly Thr Lys Val Gly Asn Val Thr Ala Lys Asp
            580                 585                 590
Pro Glu Gly Leu Asp Ile Ser Tyr Ser Leu Arg Gly Asp Thr Arg Gly
        595                 600                 605
Trp Leu Lys Ile Asp His Val Thr Gly Glu Ile Phe Ser Val Ala Pro
    610                 615                 620
Leu Asp Arg Glu Ala Gly Ser Pro Tyr Arg Val Gln Val Val Ala Thr
625                 630                 635                 640
Glu Val Gly Gly Ser Ser Leu Ser Ser Val Ser Glu Phe His Leu Ile
                645                 650                 655
Leu Met Asp Val Asn Asp Asn Pro Pro Arg Leu Ala Lys Asp Tyr Thr
            660                 665                 670
Gly Leu Phe Phe Cys His Pro Leu Ser Ala Pro Gly Ser Leu Ile Phe
        675                 680                 685
Glu Ala Thr Asp Asp Asp Gln His Leu Phe Arg Gly Pro His Phe Thr
    690                 695                 700
Phe Ser Leu Gly Ser Gly Ser Leu Gln Asn Asp Trp Glu Val Ser Lys
705                 710                 715                 720
```

```
Ile Asn Gly Thr His Ala Arg Leu Ser Thr Arg His Thr Asp Phe Glu
                725                 730                 735
Glu Arg Ala Tyr Val Val Leu Ile Arg Ile Asn Asp Gly Gly Arg Pro
            740                 745                 750
Pro Leu Glu Gly Ile Val Ser Leu Pro Val Thr Phe Cys Ser Cys Val
        755                 760                 765
Glu Gly Ser Cys Phe Arg Pro Ala Gly His Gln Thr Gly Ile Pro Thr
    770                 775                 780
Val Gly Met Ala Val Gly Ile Leu Leu Thr Thr Leu Leu Val Ile Gly
785                 790                 795                 800
Ile Ile Leu Ala Val Val Phe Ile Arg Ile Lys Lys Asp Lys Gly Lys
                805                 810                 815
Asp Asn Val Glu Ser Ala Gln Ala Ser Glu Val Lys Pro Leu Arg Ser
            820                 825                 830
```

(2) INFORMATION FOR SEQ ID NO:179:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1827 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

```
Met Ala Arg Lys Lys Phe Ser Gly Leu Glu Ile Ser Leu Ile Val Leu
 1               5                  10                  15
Phe Val Ile Val Thr Ile Ile Ala Ile Ala Leu Ile Val Val Leu Ala
            20                  25                  30
Thr Lys Thr Pro Ala Val Asp Glu Ile Ser Asp Ser Thr Ser Thr Pro
        35                  40                  45
Ala Thr Thr Arg Val Thr Thr Asn Pro Ser Asp Ser Gly Lys Cys Pro
    50                  55                  60
Asn Val Leu Asn Asp Pro Val Asn Val Arg Ile Asn Cys Ile Pro Glu
65                  70                  75                  80
Gln Phe Pro Thr Glu Gly Ile Cys Ala Gln Arg Gly Cys Cys Trp Arg
                85                  90                  95
Pro Trp Asn Asp Ser Leu Ile Pro Trp Cys Phe Phe Val Asp Asn His
            100                 105                 110
Gly Tyr Asn Val Gln Asp Met Thr Thr Thr Ser Ile Gly Val Glu Ala
        115                 120                 125
Lys Leu Asn Arg Ile Pro Ser Pro Thr Leu Phe Gly Asn Asp Ile Asn
    130                 135                 140
Ser Val Leu Phe Thr Thr Gln Asn Gln Thr Pro Asn Arg Phe Arg Phe
145                 150                 155                 160
Lys Ile Thr Asp Pro Asn Asn Arg Arg Tyr Glu Val Pro His Gln Tyr
                165                 170                 175
Val Lys Glu Phe Thr Gly Pro Thr Val Ser Asp Thr Leu Tyr Asp Val
            180                 185                 190
Lys Val Ala Gln Asn Pro Phe Ser Ile Gln Val Ile Arg Lys Ser Asn
        195                 200                 205
Gly Lys Thr Leu Phe Asp Thr Ser Ile Gly Pro Leu Val Tyr Ser Asp
    210                 215                 220
Gln Tyr Leu Gln Ile Ser Ala Arg Leu Pro Ser Asp Tyr Ile Tyr Gly
225                 230                 235                 240
Ile Gly Glu Gln Val His Lys Arg Phe Arg His Asp Leu Ser Trp Lys
                245                 250                 255
Thr Trp Pro Ile Phe Thr Arg Asp Gln Leu Pro Gly Asp Asn Asn Asn
            260                 265                 270
Asn Leu Tyr Gly His Gln Thr Phe Phe Met Cys Ile Glu Asp Thr Ser
        275                 280                 285
Gly Lys Ser Phe Gly Val Phe Leu Met Asn Ser Asn Ala Met Glu Ile
    290                 295                 300
Phe Ile Gln Pro Thr Pro Ile Val Thr Tyr Arg Val Thr Gly Gly Ile
```

```
305                 310                 315                 320
Leu Asp Phe Tyr Ile Leu Leu Gly Asp Thr Pro Glu Gln Val Val Gln
                325                 330                 335
Gln Tyr Gln Gln Leu Val Gly Leu Pro Ala Met Pro Ala Tyr Trp Asn
            340                 345                 350
Leu Gly Phe Gln Leu Ser Arg Trp Asn Tyr Lys Ser Leu Asp Val Val
        355                 360                 365
Lys Glu Val Val Arg Arg Asn Arg Glu Ala Gly Ile Pro Phe Asp Thr
    370                 375                 380
Gln Val Thr Asp Ile Asp Tyr Met Glu Asp Lys Lys Asp Phe Thr Tyr
385                 390                 395                 400
Asp Gln Val Ala Phe Asn Gly Leu Pro Gln Phe Val Gln Asp Leu His
                405                 410                 415
Asp His Gly Gln Lys Tyr Val Ile Ile Leu Asp Pro Ala Ile Ser Ile
            420                 425                 430
Gly Arg Arg Ala Asn Gly Thr Thr Tyr Ala Thr Tyr Glu Arg Gly Asn
        435                 440                 445
Thr Gln His Val Trp Ile Asn Glu Ser Asp Gly Ser Thr Pro Ile Ile
    450                 455                 460
Gly Glu Val Trp Pro Gly Leu Thr Val Tyr Pro Asp Phe Thr Asn Pro
465                 470                 475                 480
Asn Cys Ile Asp Trp Trp Ala Asn Glu Cys Ser Ile Phe His Gln Glu
                485                 490                 495
Val Gln Tyr Asp Gly Leu Trp Ile Asp Met Asn Glu Val Ser Ser Phe
            500                 505                 510
Ile Gln Gly Ser Thr Lys Gly Cys Asn Val Asn Lys Leu Asn Tyr Pro
        515                 520                 525
Pro Phe Thr Pro Asp Ile Leu Asp Lys Leu Met Tyr Ser Lys Thr Ile
    530                 535                 540
Cys Met Asp Ala Val Gln Asn Trp Gly Lys Gln Tyr Asp Val His Ser
545                 550                 555                 560
Leu Tyr Gly Tyr Ser Met Ala Ile Ala Thr Glu Gln Ala Val Gln Lys
                565                 570                 575
Val Phe Pro Asn Lys Arg Ser Phe Ile Leu Thr Arg Ser Thr Phe Ala
            580                 585                 590
Gly Ser Gly Arg His Ala Ala His Trp Leu Gly Asp Asn Thr Ala Ser
        595                 600                 605
Trp Glu Gln Met Glu Trp Ser Ile Thr Gly Met Leu Glu Phe Ser Leu
    610                 615                 620
Phe Gly Ile Pro Leu Val Gly Ala Asp Ile Cys Gly Phe Val Ala Glu
625                 630                 635                 640
Thr Thr Glu Glu Leu Cys Arg Arg Trp Met Gln Leu Gly Ala Phe Tyr
                645                 650                 655
Pro Phe Ser Arg Asn His Asn Ser Asp Gly Tyr Glu His Gln Asp Pro
            660                 665                 670
Ala Phe Phe Gly Gln Asn Ser Leu Leu Val Lys Ser Ser Arg Gln Tyr
        675                 680                 685
Leu Thr Ile Arg Tyr Thr Leu Leu Pro Phe Leu Tyr Thr Leu Phe Tyr
    690                 695                 700
Lys Ala His Val Phe Gly Glu Thr Val Ala Arg Pro Val Leu His Glu
705                 710                 715                 720
Phe Tyr Glu Asp Thr Asn Ser Trp Ile Glu Asp Thr Glu Phe Leu Trp
                725                 730                 735
Gly Pro Ala Leu Leu Ile Thr Pro Val Leu Lys Gln Gly Ala Asp Thr
            740                 745                 750
Val Ser Ala Tyr Ile Pro Asp Ala Ile Trp Tyr Asp Tyr Glu Ser Gly
        755                 760                 765
Ala Lys Arg Pro Trp Arg Lys Gln Arg Val Asp Met Tyr Leu Pro Ala
    770                 775                 780
Asp Lys Ile Gly Leu His Leu Arg Gly Gly Tyr Ile Ile Pro Ile Gln
785                 790                 795                 800
Glu Pro Asp Val Thr Thr Thr Ala Ser Arg Lys Asn Pro Leu Gly Leu
                805                 810                 815
Ile Val Ala Leu Gly Glu Asn Asn Thr Ala Lys Gly Asp Phe Phe Trp
            820                 825                 830
Asp Asp Gly Glu Thr Lys Asp Thr Ile Gln Asn Gly Asn Tyr Ile Leu
```

```
          835                 840                 845
Tyr Thr Phe Ser Val Ser Asn Asn Thr Leu Asp Ile Val Cys Thr His
    850                 855                 860
Ser Ser Tyr Gln Glu Gly Thr Thr Leu Ala Phe Gln Thr Val Lys Ile
865                 870                 875                 880
Leu Gly Leu Thr Asp Ser Val Thr Glu Val Arg Val Ala Glu Asn Asn
                885                 890                 895
Gln Pro Met Asn Ala His Ser Asn Phe Thr Tyr Asp Ala Ser Asn Gln
            900                 905                 910
Val Leu Leu Ile Ala Asp Leu Lys Leu Asn Leu Gly Arg Asn Phe Ser
        915                 920                 925
Val Gln Trp Asn Gln Ile Phe Ser Glu Asn Glu Arg Phe Asn Cys Tyr
    930                 935                 940
Pro Asp Ala Asp Leu Ala Thr Glu Gln Lys Cys Thr Gln Arg Gly Cys
945                 950                 955                 960
Val Trp Arg Thr Gly Ser Ser Leu Ser Lys Ala Pro Glu Cys Tyr Phe
                965                 970                 975
Pro Arg Gln Asp Asn Ser Tyr Ser Val Asn Ser Ala Arg Tyr Ser Ser
            980                 985                 990
Met Gly Ile Thr Ala Asp Leu Gln Leu Asn Thr Ala Asn Ala Arg Ile
        995                1000                1005
Lys Leu Pro Ser Asp Pro Ile Ser Thr Leu Arg Val Glu Val Lys Tyr
   1010                1015                1020
His Lys Asn Asp Met Leu Gln Phe Lys Ile Tyr Asp Pro Gln Lys Lys
025                1030                1035                1040
Arg Tyr Glu Val Pro Val Pro Leu Asn Ile Pro Thr Thr Pro Ile Ser
                1045                1050                1055
Thr Tyr Glu Asp Arg Leu Tyr Asp Val Glu Ile Lys Glu Asn Pro Phe
            1060                1065                1070
Gly Ile Gln Ile Arg Arg Arg Ser Ser Gly Arg Val Ile Trp Asp Ser
        1075                1080                1085
Trp Leu Pro Gly Phe Ala Phe Asn Asp Gln Phe Ile Gln Ile Ser Thr
   1090                1095                1100
Arg Leu Pro Ser Glu Tyr Ile Tyr Gly Phe Gly Glu Val Glu His Thr
105                1110                1115                1120
Ala Phe Lys Arg Asp Leu Asn Trp Asn Thr Trp Gly Met Phe Thr Arg
                1125                1130                1135
Asp Gln Pro Pro Gly Tyr Lys Leu Asn Ser Tyr Gly Phe His Pro Tyr
            1140                1145                1150
Tyr Met Ala Leu Glu Glu Glu Gly Asn Ala His Gly Val Phe Leu Leu
        1155                1160                1165
Asn Ser Asn Ala Met Asp Val Thr Phe Gln Pro Thr Pro Ala Leu Thr
   1170                1175                1180
Tyr Arg Thr Val Gly Gly Ile Leu Asp Phe Tyr Met Phe Leu Gly Pro
185                1190                1195                1200
Thr Pro Gln Val Ala Thr Lys Gln Tyr His Glu Val Ile Gly His Pro
                1205                1210                1215
Val Met Pro Ala Tyr Trp Ala Leu Gly Phe Gln Leu Cys Arg Tyr Gly
            1220                1225                1230
Tyr Ala Asn Thr Ser Glu Val Arg Glu Leu Tyr Asp Ala Met Val Ala
        1235                1240                1245
Ala Asn Ile Pro Tyr Asp Val Gln Tyr Thr Asp Ile Asp Tyr Met Glu
   1250                1255                1260
Arg Gln Leu Asp Phe Thr Ile Gly Glu Ala Phe Gln Asp Leu Pro Gln
265                1270                1275                1280
Phe Val Asp Lys Ile Arg Gly Glu Gly Met Arg Tyr Ile Ile Ile Leu
                1285                1290                1295
Asp Pro Ala Ile Ser Gly Asn Glu Thr Lys Thr Tyr Pro Ala Phe Glu
            1300                1305                1310
Arg Gly Gln Gln Asn Asp Val Phe Val Lys Trp Pro Asn Thr Asn Asp
        1315                1320                1325
Ile Cys Trp Ala Lys Val Trp Pro Asp Leu Pro Asn Ile Thr Ile Asp
   1330                1335                1340
Lys Thr Leu Thr Glu Asp Glu Ala Val Asn Ala Ser Arg Ala His Val
345                1350                1355                1360
Ala Phe Pro Asp Phe Phe Arg Thr Ser Thr Ala Glu Trp Trp Ala Arg
```

```
                1365                1370                1375
Glu Ile Val Asp Phe Tyr Asn Glu Lys Met Lys Phe Asp Gly Leu Trp
            1380                1385                1390
Ile Asp Met Asn Glu Pro Ser Ser Phe Val Asn Gly Thr Thr Thr Asn
        1395                1400                1405
Gln Cys Arg Asn Asp Glu Leu Asn Tyr Pro Pro Tyr Phe Pro Glu Leu
    1410                1415                1420
Thr Lys Arg Thr Asp Gly Leu His Phe Arg Thr Ile Cys Met Glu Ala
425                 1430                1435                1440
Glu Gln Ile Leu Ser Asp Gly Thr Ser Val Leu His Tyr Asp Val His
                1445                1450                1455
Asn Leu Tyr Gly Trp Ser Gln Met Lys Pro Thr His Asp Ala Leu Gln
            1460                1465                1470
Lys Thr Thr Gly Lys Arg Gly Ile Val Ile Ser Arg Ser Thr Tyr Pro
        1475                1480                1485
Thr Ser Gly Arg Trp Gly Gly His Trp Leu Gly Asp Asn Tyr Ala Arg
    1490                1495                1500
Trp Asp Asn Met Asp Lys Ser Ile Ile Gly Met Met Glu Phe Ser Leu
505                 1510                1515                1520
Phe Gly Ile Ser Tyr Thr Gly Ala Asp Ile Cys Gly Phe Phe Asn Asn
                1525                1530                1535
Ser Glu Tyr His Leu Cys Thr Arg Trp Met Gln Leu Gly Ala Phe Tyr
            1540                1545                1550
Pro Tyr Ser Arg Asn His Asn Ile Ala Asn Thr Arg Arg Gln Asp Pro
        1555                1560                1565
Ala Ser Trp Asn Glu Thr Phe Ala Glu Met Ser Arg Asn Ile Leu Asn
    1570                1575                1580
Ile Arg Tyr Thr Leu Leu Pro Tyr Phe Tyr Thr Gln Met His Glu Ile
585                 1590                1595                1600
His Ala Asn Gly Gly Thr Val Ile Arg Pro Leu Leu His Glu Phe Phe
                1605                1610                1615
Asp Glu Lys Pro Thr Trp Asp Ile Phe Lys Gln Phe Leu Trp Gly Pro
            1620                1625                1630
Ala Phe Met Val Thr Pro Val Leu Glu Pro Tyr Val Gln Thr Val Asn
        1635                1640                1645
Ala Tyr Val Pro Asn Ala Arg Trp Phe Asp Tyr His Thr Gly Lys Asp
    1650                1655                1660
Ile Gly Val Arg Gly Gln Phe Gln Thr Phe Asn Ala Ser Tyr Asp Thr
665                 1670                1675                1680
Ile Asn Leu His Val Arg Gly Gly His Ile Leu Pro Cys Gln Glu Pro
                1685                1690                1695
Ala Gln Asn Thr Phe Tyr Ser Arg Gln Lys His Met Lys Leu Ile Val
            1700                1705                1710
Ala Ala Asp Asp Asn Gln Met Ala Gln Gly Ser Leu Phe Trp Asp Asp
        1715                1720                1725
Gly Glu Ser Ile Asp Thr Tyr Glu Arg Asp Leu Tyr Leu Ser Val Gln
    1730                1735                1740
Phe Asn Leu Asn Gln Thr Thr Leu Thr Ser Thr Ile Leu Lys Arg Gly
745                 1750                1755                1760
Tyr Ile Asn Lys Ser Glu Thr Arg Leu Gly Ser Leu His Val Trp Gly
                1765                1770                1775
Lys Gly Thr Thr Pro Val Asn Ala Val Thr Leu Thr Tyr Asn Gly Asn
            1780                1785                1790
Lys Asn Ser Leu Pro Phe Asn Glu Asp Thr Thr Asn Met Ile Leu Arg
        1795                1800                1805
Ile Asp Leu Thr Thr His Asn Val Thr Leu Glu Glu Pro Ile Glu Ile
    1810                1815                1820
Asn Trp Ser
825
```

(2) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2284 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

```
CTT TTC ACA CTC CCT GGA ACT CCT ATA ACT TAC TAT GGA GAA GAA ATT    1496
Leu Phe Thr Leu Pro Gly Thr Pro Ile Thr Tyr Tyr Gly Glu Glu Ile
    470                 475                 480

GGA ATG GGA AAT ATT GTA GCC GCA AAT CTC AAT GAA AGC TAT GAT ATT    1544
Gly Met Gly Asn Ile Val Ala Ala Asn Leu Asn Glu Ser Tyr Asp Ile
485                 490                 495                 500

AAT ACC CTT CGC TCA AAG TCA CCA ATG CAG TGG GAC AAT AGT TCA AAT    1592
Asn Thr Leu Arg Ser Lys Ser Pro Met Gln Trp Asp Asn Ser Ser Asn
                505                 510                 515

GCT GGT TTT TCT GAA GCT AGT AAC ACC TGG TTA CCT ACC AAT TCA GAT    1640
Ala Gly Phe Ser Glu Ala Ser Asn Thr Trp Leu Pro Thr Asn Ser Asp
            520                 525                 530

TAC CAC ACT GTG AAT GTT GAT GTC CAA AAG ACT CAG CCC AGA TCG GCT    1688
Tyr His Thr Val Asn Val Asp Val Gln Lys Thr Gln Pro Arg Ser Ala
        535                 540                 545

TTG AAG TTA TAT CAA GAT TTA AGT CTA CTT CAT GCC AAT GAG CTA CTC    1736
Leu Lys Leu Tyr Gln Asp Leu Ser Leu Leu His Ala Asn Glu Leu Leu
    550                 555                 560

CTC AAC AGG GGC TGG TTT TGC CAT TTG AGG AAT GAC AGC CAC TAT GTT    1784
Leu Asn Arg Gly Trp Phe Cys His Leu Arg Asn Asp Ser His Tyr Val
565                 570                 575                 580

GTG TAC ACA AGA GAG CTG GAT GGC ATC GAC AGA ATC TTT ATC GTG GTT    1832
Val Tyr Thr Arg Glu Leu Asp Gly Ile Asp Arg Ile Phe Ile Val Val
                585                 590                 595

CTG AAT TTT GGA GAA TCA ACA CTG TTA AAT CTA CAT AAT ATG ATT TCG    1880
Leu Asn Phe Gly Glu Ser Thr Leu Leu Asn Leu His Asn Met Ile Ser
            600                 605                 610

GGC CTT CCC GCT AAA ATA AGA ATA AGG TTA AGT ACC AAT TCT GCC GAC    1928
Gly Leu Pro Ala Lys Ile Arg Ile Arg Leu Ser Thr Asn Ser Ala Asp
        615                 620                 625

AAA GGC AGT AAA GTT GAT ACA AGT GGC ATT TTT CTG GAC AAG GGA GAG    1976
Lys Gly Ser Lys Val Asp Thr Ser Gly Ile Phe Leu Asp Lys Gly Glu
    630                 635                 640

GGA CTC ATC TTT GAA CAC AAC ACG AAG AAT CTC CTT CAT CGC CAA ACA    2024
Gly Leu Ile Phe Glu His Asn Thr Lys Asn Leu Leu His Arg Gln Thr
645                 650                 655                 660

GCT TTC AGA GAT AGA TGC TTT GTT TCC AAT CGA GCA TGC TAT TCC AGT    2072
Ala Phe Arg Asp Arg Cys Phe Val Ser Asn Arg Ala Cys Tyr Ser Ser
                665                 670                 675

GTA CTG AAC ATA CTG TAT ACC TCG TGT TAGGCACCTT TATGAAGAGA TGAAGAC  2126
Val Leu Asn Ile Leu Tyr Thr Ser Cys
            680                 685

ACTGGCATTT CAGTGGGATT GTAAGCATTT GTAATAGCTT CATGTACAGC ATGCTGCTTG   2186
GTGAACAATC ATTAATTCTT CGATATTTCT GTAGCTTGAA TGTAACCGCT TTAAGAAAGG   2246
TTCTCAAATG TTTTGAAAAA AATAAAATGT TTAAAAGT                           2284
```

(2) INFORMATION FOR SEQ ID NO:181:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 685 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

```
Met Ala Glu Asp Lys Ser Lys Arg Asp Ser Ile Glu Met Ser Met Lys
 1               5                  10                  15
Gly Cys Gln Thr Asn Asn Gly Phe Val His Asn Glu Asp Ile Leu Glu
            20                  25                  30
Gln Thr Pro Asp Pro Gly Ser Ser Thr Asp Asn Leu Lys His Ser Thr
        35                  40                  45
Arg Gly Ile Leu Gly Ser Gln Glu Pro Asp Phe Lys Gly Val Gln Pro
    50                  55                  60
Tyr Ala Gly Met Pro Lys Glu Val Leu Phe Gln Phe Ser Gly Gln Ala
65                  70                  75                  80
Arg Tyr Arg Ile Pro Arg Glu Ile Leu Phe Trp Leu Thr Val Ala Ser
                85                  90                  95
Val Leu Val Leu Ile Ala Ala Thr Ile Ala Ile Ile Ala Leu Ser Pro
            100                 105                 110
Lys Cys Leu Asp Trp Trp Gln Glu Gly Pro Met Tyr Gln Ile Tyr Pro
        115                 120                 125
Arg Ser Phe Lys Asp Ser Asn Lys Asp Gly Asn Gly Asp Leu Lys Gly
    130                 135                 140
Ile Gln Asp Lys Leu Asp Tyr Ile Thr Ala Leu Asn Ile Lys Thr Val
145                 150                 155                 160
Trp Ile Thr Ser Phe Tyr Lys Ser Ser Leu Lys Asp Phe Arg Tyr Gly
                165                 170                 175
Val Glu Asp Phe Arg Glu Val Asp Pro Ile Phe Gly Thr Met Glu Asp
            180                 185                 190
Phe Glu Asn Leu Val Ala Ala Ile His Asp Lys Gly Leu Lys Leu Ile
        195                 200                 205
Ile Asp Phe Ile Pro Asn His Thr Ser Asp Lys His Ile Trp Phe Gln
    210                 215                 220
Leu Ser Arg Thr Arg Thr Gly Lys Tyr Thr Asp Tyr Tyr Ile Trp His
225                 230                 235                 240
Asp Cys Thr His Glu Asn Gly Lys Thr Ile Pro Pro Asn Asn Trp Leu
                245                 250                 255
Ser Val Tyr Gly Asn Ser Ser Trp His Phe Asp Glu Val Arg Asn Gln
            260                 265                 270
Cys Tyr Phe His Gln Phe Met Lys Glu Gln Pro Asp Leu Asn Phe Arg
        275                 280                 285
Asn Pro Asp Val Gln Glu Glu Ile Lys Glu Ile Leu Arg Phe Trp Leu
    290                 295                 300
Thr Lys Gly Val Asp Gly Phe Ser Leu Asp Ala Val Lys Phe Leu Leu
305                 310                 315                 320
Glu Ala Lys His Leu Arg Asp Glu Ile Gln Val Asn Lys Thr Gln Ile
                325                 330                 335
Pro Asp Thr Val Thr Gln Tyr Ser Glu Leu Tyr His Asp Phe Thr Thr
            340                 345                 350
Thr Gln Val Gly Met His Asp Ile Val Arg Ser Phe Arg Gln Thr Met
        355                 360                 365
Asp Gln Tyr Ser Thr Glu Pro Gly Arg Tyr Arg Phe Met Gly Thr Glu
    370                 375                 380
Ala Tyr Ala Glu Ser Ile Asp Arg Thr Val Met Tyr Tyr Gly Leu Pro
385                 390                 395                 400
Phe Ile Gln Glu Ala Asp Phe Pro Phe Asn Asn Tyr Leu Ser Met Leu
                405                 410                 415
Asp Thr Val Ser Gly Asn Ser Val Tyr Glu Val Ile Thr Ser Trp Met
            420                 425                 430
Glu Asn Met Pro Glu Gly Lys Trp Pro Asn Trp Met Ile Gly Gly Pro
        435                 440                 445
Asp Ser Ser Arg Leu Thr Ser Arg Leu Gly Asn Gln Tyr Val Asn Val
    450                 455                 460
Met Asn Met Leu Leu Phe Thr Leu Pro Gly Thr Pro Ile Thr Tyr Tyr
465                 470                 475                 480
```

```
Gly Glu Glu Ile Gly Met Gly Asn Ile Val Ala Ala Asn Leu Asn Glu
                485                 490                 495
Ser Tyr Asp Ile Asn Thr Leu Arg Ser Lys Ser Pro Met Gln Trp Asp
            500                 505                 510
Asn Ser Ser Asn Ala Gly Phe Ser Glu Ala Ser Asn Thr Trp Leu Pro
        515                 520                 525
Thr Asn Ser Asp Tyr His Thr Val Asn Val Asp Val Gln Lys Thr Gln
    530                 535                 540
Pro Arg Ser Ala Leu Lys Leu Tyr Gln Asp Leu Ser Leu Leu His Ala
545                 550                 555                 560
Asn Glu Leu Leu Leu Asn Arg Gly Trp Phe Cys His Leu Arg Asn Asp
                565                 570                 575
Ser His Tyr Val Val Tyr Thr Arg Glu Leu Asp Gly Ile Asp Arg Ile
            580                 585                 590
Phe Ile Val Val Leu Asn Phe Gly Glu Ser Thr Leu Leu Asn Leu His
        595                 600                 605
Asn Met Ile Ser Gly Leu Pro Ala Lys Ile Arg Ile Arg Leu Ser Thr
    610                 615                 620
Asn Ser Ala Asp Lys Gly Ser Lys Val Asp Thr Ser Gly Ile Phe Leu
625                 630                 635                 640
Asp Lys Gly Glu Gly Leu Ile Phe Glu His Asn Thr Lys Asn Leu Leu
                645                 650                 655
His Arg Gln Thr Ala Phe Arg Asp Arg Cys Phe Val Ser Asn Arg Ala
            660                 665                 670
Cys Tyr Ser Ser Val Leu Asn Ile Leu Tyr Thr Ser Cys
        675                 680                 685
```

(2) INFORMATION FOR SEQ ID NO:182:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

```
Leu Val Pro Arg Gly Ser Pro Gly Ile Pro Gly Ser Arg Val Gly Gln
 1               5                  10                  15
Cys Thr Asp Ser Asp Val Arg Arg Pro Trp Ala Arg Ser Cys Ala His
            20                  25                  30
Gln Gly Cys Gly Ala Gly Thr Arg Asn Ser His Gly Cys Ile Thr Arg
        35                  40                  45
Pro Leu Arg Gln Ala Ser
    50
```

(2) INFORMATION FOR SEQ ID NO:183:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

```
Ser Ala Arg Asp Ser Gly Pro Ala Glu Asp Gly Ser Arg Ala Val Arg
 1               5                  10                  15
Leu Asn Gly
```

(2) INFORMATION FOR SEQ ID NO:184:

WHAT IS CLAIMED IS:

1. A purified protein which specifically binds to a gastro-intestinal tract receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI.

2. A protein which binds specifically to a gastro-intestinal transport receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI, in which the protein comprises an amino acid sequence selected from the group consisting of SEQ ID NOS:1-55 or a binding portion thereof.

3. A protein which binds specifically to a gastro-intestinal transport receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI, in which the amino acid sequence of the protein is selected from the group consisting of SEQ ID NOS:1-55, or a binding portion thereof.

4. The protein of claim 2 which comprises the amino acid sequence substantially as set forth in: SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 30, SEQ ID NO: 43, SEQ ID NO: 46, or SEQ ID NO: 52, or a binding portion thereof.

5. The protein of claim 3, the amino acid sequence of which consists of the amino acid sequence substantially as set forth in: SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 30, SEQ ID NO: 43, SEQ ID NO: 46, or SEQ ID NO: 52, or a binding portion thereof.

6. A protein of not more than 50 amino acids in length which specifically binds to a gastro-intestinal transport receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI, in which the protein includes, positioned anywhere along its sequence, the contiguous amino

acid sequence of: Xaa_1 Thr Xaa_2 Xaa_3 Ser Xaa_4 Xaa_5 Xaa_6 Asn Xaa_7 Arg (SEQ ID NO:253), where Xaa_1 is Ser or Thr; Xaa_2 is Arg or Lys; Xaa_3 is Lys or Arg; Xaa_4 is Ser or Leu; Xaa_5 is Arg, Ile, Val, or Ser; Xaa_6 is Ser, Tyr, Phe, or His; and Xaa_7 is Pro, His or Arg.

7. The protein of claim 6 which is not more than 40 amino acids in length.

8. The protein of claim 6 which is not more than 30 amino acids in length.

9. The protein of claim 6 which is not more than 20 amino acids in length.

10. A protein of not more than 50 amino acids in length which specifically binds to a gastro-intestinal transport receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI, in which the protein includes, positioned anywhere along its sequence, the contiguous amino acid sequence of: Asp Xaa_1 Asp Xaa_2 Arg Arg Xaa_3 Xaa_4 (SEQ ID NO:254) where Xaa_1 is Ser, Ala, or Gly; Xaa_2 is Val or Gln; Xaa_3 is Pro, Gly, or Ser; and Xaa_4 is Trp or Tyr.

11. The protein of claim 10 which is not more than 40 amino acids in length.

12. The protein of claim 10 which is not more than 30 amino acids in length.

13. The protein of claim 10 which is not more than 20 amino acids in length.

14. A protein of not more than 50 amino acids in length which specifically binds to a gastro-intestinal transport receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI, in which the protein includes,

positioned anywhere along its sequence, the contiguous amino acid sequence of: Val Arg Ser Gly Cys Gly Xaa_1 Xaa_2 Ser Ser (SEQ ID NO:255), where Xaa_1 is Ala or Phe; and Xaa_2 is Arg or His.

15. The protein of claim 14 which is not more than 40 amino acids in length.

16. The protein of claim 14 which is not more than 30 amino acids in length.

17. The protein of claim 14 which is not more than 20 amino acids in length.

18. A protein of not more than 50 amino acids in length which specifically binds to a gastro-intestinal transport receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI, in which the protein includes, positioned anywhere along its sequence, the contiguous amino acid sequence of: NTRKSSRSNPR (SEQ ID NO:256) or STKRSLIYNHR (SEQ ID NO:257) or STGRKVFNRR (SEQ ID NO:258) or TNAKHSSHNRR (SEQ ID NO:259).

19. A protein of not more than 50 amino acids in length which specifically binds to a gastro-intestinal transport receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI, in which the protein includes, positioned anywhere along its sequence, the contiguous amino acid sequence of: DSDVRRPW (SEQ ID NO:260) or AADQRRGW (SEQ ID NO:261) or DGRGGRSY (SEQ ID NO:262).

20. A protein of not more than 50 amino acids in length which specifically binds to a gastro-intestinal transport receptor selected from the group consisting of HPT1, hPEPT1, D2H, and hSI, in which the protein includes, positioned anywhere along its sequence, the contiguous amino

acid sequence of: RVRS (SEQ ID NO:263) or SVRSGCGFRGSS (SEQ ID NO:264) or SVRGGCGAHSS (SEQ ID NO:265).

21. The protein of claim 1, 2, 3, 6, 10, 14, 18, 19, or 20 which is purified.

22. A composition comprising the protein of claim 1, 2, 3, 6, 10, 14, 18, 19, or 20, bound to a material comprising an active agent, said active agent being of value in the treatment of a mammalian disease or disorder.

23. The composition of claim 22 in which the active agent is a drug.

24. The composition of claim 22 in which the material is a particle containing the active agent.

25. The composition of claim 22 in which the material is a slow-release device containing the drug.

26. The composition of claim 22 in which the protein is covalently or noncovalently bound to the material.

27. A composition comprising a chimeric protein bound to a material comprising an active agent, in which the chimeric protein comprises a sequence selected from the group consisting of SEQ ID NOS:1-55 or a binding portion thereof fused via a covalent bond to an amino acid sequence of a second protein, in which the active agent is of value in the treatment of a mammalian disease or disorder.

28. A composition comprising the protein of claim 1, 2, 3, 6, 10, 14, 18, 19, or 20 covalently bound to a particle containing a drug.

29. A composition comprising the protein of claim 1, 2, 3, 6, 10, 14, 18, 19, or 20 covalently bound to a drug.